DOCUMENT RESUME

ED 058 294 TM 001 009

AUTHOR Blai, Boris, Jr.
TITLE Faculty Effectiveness--A 'Pilot' Study of Student Evaluation at Harcum Junior College.
INSTITUTION Harcum Junior Coll., Bryn Mawr, Pa.
REPORT NO IRR-71-37
PUB DATE Dec 71
NOTE 5p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Check Lists; College Instruction; *Effective Teaching; Evaluation Criteria; Pilot Projects; *Rating Scales; *Student Opinion; Summative Evaluation; Teacher Behavior; *Teacher Evaluation; Teacher Rating; *Teaching Quality
IDENTIFIERS *Faculty Evaluation Check List; Harcum Junior College

ABSTRACT

Approximately 42 percent of the Harcum Junior College student body participated in a study of faculty effectiveness. Results of the Faculty Evaluation Check List, which consisted of 10 Likert-type items and two open-ended questions, indicate that students have high regard for their professors. A copy of the instrument is included. (AG)

HARCUM JUNIOR COLLEGE
BRYN MAWR, PENNA. 19010



Office of Research
IRR 71-37

# Faculty Effectiveness-A 'Pilot' Study of Student Evaluation at Harcum Junior College

1. "A fundamental assumption relating to evaluation is that such evaluation of faculty instruction through some form of rating, or measurement device, is of value."(Blai, p. 153) The purpose of this particular Institutional Research Report (IRR 71-37) is to provide the faculty and administration with individually unidentifiable student opinions on various aspects of faculty teaching performance, and to present a collective 'picture' of faculty effectiveness as viewed by a very substant ial-sized sample of the Harcum student body. "The items ... in the Faculty Evaluation Check-list are offered solely because they appear, in my judgment, to have "face validity" as items reflecting teacher behavior which can contribute to the learning of students. They are not offered as items selected through "item analysis"; "factor analysis," or other sophisticated statistical techniques. They simply seem to 'make sense.'" (Blai, p. 155)

2. "The growing body of research literature on the use of student ratings of teachers indicates that they can and do make reasonable, accurate ratings (Simpson, 1966). As Howsan (1960) reports: their ratings tend to agree with each other, and the teachers who are rated best by the pupils tend to obtain the highest pupil gains. Pupil ratings often do not agree with the ratings by principals, supervisors or teachers. (This has not been considered as an indication of weakness in student ratings, however, since ratings by supervisors and peers have not been shown to agree with pupil-gained measures or to be satisfactory in other ways.) Teachers have indicated their belief that pupil ratings, as obtained in research studies, are both fair and accurate." (Blai, p. 154)

3. The very sizeable student-population samples, selected completely at random, who participated in this pilot evaluation during academic year 1970-71, is reflected in the following tabulation:

| Division | | Average size of class samples |
|---|---|---|
| Business | = | 43% |
| Fine Arts | = | 42% |
| Education | = | 40% |
| English & Language | = | 40% |
| Science & Math | = | 43% |
| Social Sciences | = | 47% |

Clearly then, the comments which follow reflect the concensus of views of a substantial number (approximately 270) of the Harcum student population of 630.

4. A useful method of comparing experimentally obtained results with those to be expected theoretically on the hypothesis of equal probability (the so-called "null hypothesis) is the Chi-square ($X^2$) test. Application of this formula permits one to determine if statistically significant differences exist between 'theoretical' and 'obtained' results. The more closely the observed results approximate the expected, the smaller the chi-square, and the closer the agreement between observed data and the hypothesis being tested. Contrariwise, the larger the chi-square, the greater the probability of a real divergence of experimentally observed from expected results. The equation for chi-square is:-

$$X^2 = \Sigma \left[ \frac{(fo-fe)^2}{fe} \right]$$

in which:- fo = frequency of occurrence of 'observed' facts
fe = "expected" theoretical frequency of occurrence
$\Sigma$ = sum of

5. Theoretically, an equal number of faculty (20%), through chance variation alone (equal probability) would be 'rated' in each of the five adjective categories - Superior, Above Average, Average, Below Average, Poor. This assumption would be applicable for each of the 10 items included in the Faculty Evaluation Check List utilized. The "observed" ratings, in contrast to the theoretical "expected" ratings, are summarized in the following tabulation. All percentages reported are rounded off to the nearest whole number.

Faculty Evaluation Check List Ratings

| Item | Superior | Above Aver. | Aver. | Below Aver. | Poor |
|---|---|---|---|---|---|
| 1. Instructor's preparation? | 48% | 29% | 18% | 3% | 2% |
| 2. Fairness? | 38% | 32% | 21% | 7% | 2% |
| 3. Knowledge of subject? | 64% | 24% | 8% | 3% | 1% |
| 4. Lectures or discussions follow organized pattern? | 40% | 29% | 21% | 5% | 5% |
| 5. Clarity of assignments as to what is expected of students? | 37% | 27% | 23% | 5% | 8% |
| 6. A patterned timetable worked out to fit the course to number of class meetings? | 42% | 27% | 21% | 7% | 3% |
| 7. Skill in communication clarity of expression and explanation? | 43% | 28% | 19% | 6% | 4% |
| 8. Helpfulness to individual students? | 43% | 28% | 18% | 7% | 4% |
| 9. Emotional stability (even temperament, patience, and sense of humor)? | 54% | 25% | 13% | 4% | 4% |
| 10. Enthusiasm, (ability to "sell" students on desirability of their working to capacity? | 49% | 25% | 18% | 5% | 3% |

6. Were any of these differences between 'theoretical' and 'observed' frequencies statistically significant? Application of the Chi-square formula to the data yielded the following results:

Item #1 - Instructor's preparation:- $X^2 = 63.2$ with a probability or "P" = less than.0001. In general, one may safely discard a null hypothesis of "equal probability" whenever "P" is .05 or less. Clearly then, this distribution of ratings may be marked "Significant" at less than the .0001 level (i.e. less than 1 chance in 10,000 of this distribution occuring through sampling fluctuations or 'chance' alone). Therefore, the null hypothesis is rejected and it is concluded that this substantial sample of Harcum students strongly favor the teaching effectiveness of the Harcum faculty as reflected in the item - "Instructor's preparation for each class (lab session."

7. The Chi-square values for each of the other 9 items in the Faculty Evaluation Check-List were also found to be "Significant" at less than the .0001 level. Therefore, as might be anticipated, the distribution of the total 2187 student evaluations among the 5-item adjective scale of Superior (46%), Average (27%), Average (18%), Below Average (5%), and Poor (4%) was also found to be statistically significant; to a very high level of confidence. It is therefore concluded, with reason, that the Harcum faculty, as a group, are by all odds far and above average as teachers.

8. This conclusion is further supported by the write-in comments of "teacher strong points" which were offered by 60% of the 270 student-respondents. Without itemizing each of the 162 statements the theme clearly apparent throughout was the fact that Harcum faculty are student-oriented; helpful and patient with individuals and enthusiastic!

9. Equally gratifying is the very short-list of "shortcomings" - the write-in comments of "teacher weak points" which consisted of only 55 items (24%). In the main, these spoke of limited or non-existent opportunities for class discussion; lecturing too rapidly; not being sufficiently explicit; and expecting too much from students.

10. And finally, it was particularly noteworthy that in the space provided on the Check-List for writing in "Teacher's weak points", some 41 students wrote - "none"!

11. This was only a limited-scope study, the project's objective being to explore insufficiencies and problems. Hopefully this has been achieved, and it therefore augers well for the possible future full administration of some type of "composite" (Blai, 1971) faculty evaluation plan.

Boris Blai, Jr. Ed.D.
Director of Research

December 1971

Note:- Special thanks are due Ms Margaret Flagg, Class of 1972, who processed and summarized the student evaluations - the basic data in this pilot study.

## Faculty Evaluation Check List

| Items | Superior 5 | Above Average 4 | Average 3 | Below Average 2 | Poor 1 |
|---|---|---|---|---|---|
| 1. Instructor's preparation for each class (lab.) session? | | | | | |
| 2. Fairness? | | | | | |
| 3. Knowledge of subject? | | | | | |
| 4. Lectures or discussion follow organized pattern? | | | | | |
| 5. Clarity of assignments as to what is expected of students? | | | | | |
| 6. A patterned timetable worked out to fit the course to number of class meetings? | | | | | |
| 7. Skill in communication clarity of expression and explanation? | | | | | |
| 8. Helpfulness to individual students? | | | | | |
| 9. Emotional stability, (even temperament, patience, and sense of humor)? | | | | | |
| 10. Enthusiasm, (ability to "sell" students on desirability of their working to capacity? | | | | | |

Teacher's strong points _____

_____

Teacher's weak points _____

_____

Superior = Exceptional performance
Above Average = Greater than satisfactory
Average = Satisfactory
Below Average = Minimal acceptable performance
Poor = Non-acceptable performance

Appendix A

## References


Blai, Boris, Jr. Evaluation of Faculty Instruction - Scientia Paedagogica Experimentalis. (Belgium) Vol. V, No. 2 1968. pp 153-159

Howsam, R.B. Who's A Good Teacher? (1705 Merchinson Drive, Burlingame California, Joint Committee on Personnel Procedures, 1960)

Simpson, R. H. and Seidman, J. M Student Evaluation of Teaching and Learning (Washington, D.C. The American Association of Colleges For Teacher Evaluation, 1962)

Simpson, R.B. Teacher Self-Evaluation (The Macmillan Co., N.Y. 1966)